CLANS, GUILDS, AND MARKETS: APPRENTICESHIP INSTITUTIONS AND GROWTH IN THE PRE-INDUSTRIAL ECONOMY
Online Appendix

David de la Croix
Matthias Doepke
Joel Mokyr

June 22, 2017

A Extension with Farmers

In this section, we sketch how the model can be extended by including farm labor as a separate input in production. This extension addresses the concern that in the pre-industrial era, most people were engaged in food production, whereas craftspeople made up a smaller fraction of the population.

The economy produces two goods, food F and manufactured goods M. Manufactured goods are produced under constant returns with effective craftsmen's labor L as the only input:

$$M = L.$$

Food is produced using a Cobb-Douglas technology that uses land X and farm labor N_f:

$$F = (N_f)^{\frac{1-\alpha-\beta}{1-\beta}} X^{\frac{\alpha}{1-\beta}}$$

with α, $\beta \in (0,1)$. An individual has Cobb-Douglas utility over consumption of food f and manufactured goods m, and total utility (including the altruistic component) is:

$$u(c, I') = m^{\beta} f^{1-\beta} + \gamma I'. \tag{1}$$

The setup is equivalent to one with a single consumption good c (a composite of food and manufactured goods) as in (1) that is produced with the aggregate production function:

$$Y = (N_f)^{1-\alpha-\beta} L^{\beta} X^{\alpha}.$$

The total amount of land is normalized to one, $X = 1$, and land is owned by farmers.

There are now three aggregate state variables: N_f (population of farmers), N_m (population of craftsmen), and k. Let $N = N_m + N_f$ be the total number of adults. Farmers and craftsmen have the same survival rate and there is no intergenerational mobility across occupations. Hence, the laws of motion for population are:

$$N_m' = n\, N_m, \quad N_f' = n\, N_f, \quad N' = N_m' + N_f' = n\, N.$$

As a consequence, the share of both groups in the total population is constant. We define ρ as

$$\rho = \frac{N_m}{N}.$$

The assumptions of equal population growth and no occupational mobility are made for simplicity. In reality, it is well known that in pre-industrial times cities (where craftsmen were concentrated) experienced much higher mortality than the countryside, so that there was net migration into cities. Allowing for such rural-urban migration could be accommodated in our framework and would leave the main results intact, but would come at the cost of complicating the analysis. Given that our focus is on knowledge transmission rather than urbanization, we abstract from such features here.

Given those two changes to the specification, the rest of analysis carries on. Two new parameters are involved. The market equilibrium condition (18) becomes:

$$\delta'(a^{\mathbf{M}}) - \kappa = \frac{\gamma\theta\beta}{\rho} \frac{1}{a^{\mathbf{M}}/n^{\mathbf{M}} + \nu} y^{\mathbf{M}}.$$

The extension with farmers clarifies that the Malthusian constraint matters even if there are constant returns to scale in the craftsmen's sector. Along a balanced growth path, manufacturing output grows faster than food output. This depresses the relative price of manufactured goods (produced by craftsmen). Specifically, in a given period the relative price of manufactured goods is given by:

$$\frac{p_M}{p_F} = \frac{\beta}{1-\beta} \frac{(N_f)^{\frac{1-\alpha-\beta}{1-\beta}} X^{\frac{\alpha}{1-\beta}}}{L}.$$

The declining relative price of the output of craftsmen implies that their effective income (in terms of the composite consumption good) is constant in the balanced growth path even though their output is growing.

It also becomes easier to analyze the role of parameter α, as now this only plays a role in the Malthusian constraint. A low α corresponds to labor-intensive agriculture. In this case returns to population size decrease at a lower rate, and hence an increase in productivity growth leads to a larger shift in population growth and income per capita compared to the case of a large land share.[1] It modifies the incentives to move to another mode of organizing apprenticeship, as detailed in Proposition 1.

Proposition 1 (**Effect of Labor-Intensive Agriculture).**
If Agriculture is labor intensive (low α), the long-term gains in growth from moving from family to market, $g^{\mathbf{M}} - g^{\mathbf{F}}$, and from clan to market, $g^{\mathbf{M}} - g^{\mathbf{C}}$, are reduced. But the gain from moving from family to clan, $g^{\mathbf{C}} - g^{\mathbf{F}}$, is increased.

Figure I shows the result graphically. An implication of this proposition is that a country with labor intensive agriculture has more incentives to adopt the clan than the family.

B Extension with the "Leonardo da Vinci" Assumption

In this section, we allow for the possibility that advanced techniques may be "ahead of their time," in the sense that their full productive value can only be realized at a higher state of aggregate knowledge. For example, Leonardo da Vinci invented a number of machines and devices that could be successfully

1. Vollrath (2011) argues that agriculture in pre-industrial China was more labor intensive compared to Europe (due to the possibility of multiple crops per year, wet-field rice production etc.), and that this accounts for part of observed differences in living standards.

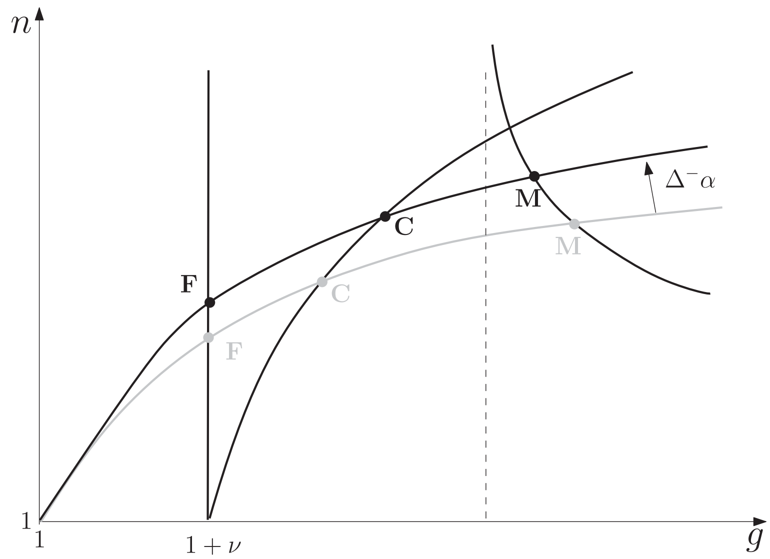

Figure I: Effect of labor intensive agriculture

built only centuries later. To capture this feature in a simple way, we assume that each craftsman has a potential output which is linked to own knowledge h_i, but that the craftsman may be constrained by the average state of knowledge in the economy. Specifically, the potential output $\bar{q}_i$ of a craftsman with knowledge h_i is given by:

$$\bar{q}_i = h_i^{-\theta}. \tag{2}$$

The actual output q_i of a craftsman cannot exceed the average potential output $\bar{q} = \mathbb{E}(\bar{q}_i)$ in the economy, so that:

$$q_i = \min\{\bar{q}_i, \bar{q}\}. \tag{3}$$

The assumption that individual productivity is constrained by aggregate knowledge is not essential to any of our results. Still, without this assumption in any period there would be some masters with arbitrarily high output, which is implausible. With the constraint, a good part of the knowledge in the economy is latent knowledge that will unfold its full potential only in later generations. This closely relates to Mokyr (2002)'s argument that the growth of productivity is constrained by the epistemic base on which a technique rests. The more people using a technique understand the science behind it, the broader the base. According to Mokyr (2002), this basis was very narrow in pre-modern Europe and gradually became wider.

Figure II illustrates the distribution of actual output q_i among craftsmen for two values of knowledge k, where the dashed line corresponds to a higher state of knowledge. The kinks in the distribution functions represent the points above which potential productivity is constrained by the average knowledge in the society. Because of the specific shape of the exponential distribution, the share of craftsmen who are constrained by the average knowledge is constant (and given by $1/e$).

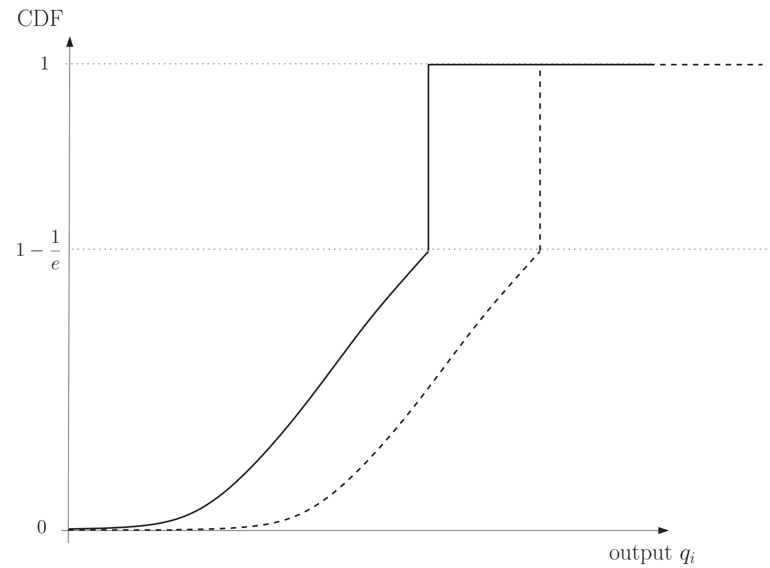

FIGURE II: DISTRIBUTION FUNCTION OF PRODUCTIVITY AT TIME t (SOLID LINE) AND $t' > t$ (DASHED LINE)

Given the exponential distribution for h_i and (2), potential output $\bar{q}_i$ follows a Fréchet distribution with scale parameter k^θ and shape parameter $1/\theta$.

We can now express the total supply of effective labor by craftsmen as a function of state variables. The average potential output across craftsmen is given by:

$$\bar{q} = \int_0^\infty h_i^{-\theta} \left(k \exp(-kh_i)\right) dh_i = \int_0^\infty k^\theta (kh_i)^{-\theta} \exp(-kh_i)\, kdh_i = k^\theta \Gamma(1-\theta),$$

where $\Gamma(t) = \int_0^\infty x^{t-1} \exp(-x)dx$ is the Euler gamma function. Given (3), the actual output q_i of a craftsman is:

$$q_i = \min\{\bar{q}_i, \bar{q}\} = \min\left[h_i^{-\theta}, k^\theta \Gamma(1-\theta)\right].$$

The threshold for h_i below which craftsmen are constrained by average knowledge is given by:

$$\hat{h} = k^{-1} \Gamma(1-\theta)^{-1/\theta}.$$

The expected supply of output of a given craftsman is:

$$\begin{aligned}\mathbb{E}(q_i) &= \int_0^{\hat{h}} k^\theta \Gamma(1-\theta) k \exp[-k h_i] dh_i + \int_{\hat{h}}^{\infty} h_i^{-\theta} k \exp[-k h_i] dh_i \\ &= k^\theta \Lambda.\end{aligned}$$

Here Λ is a constant given by:

$$\Lambda = \Gamma(1-\theta) + \exp[-\Gamma(1-\theta)^{-1/\theta}]\theta\Gamma(-\theta) + \Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta}),$$

and $\Gamma(t,s) = \int_s^\infty x^{t-1} \exp(-x) dx$ is the incomplete gamma function. The total supply of craftsmen's labor L in efficiency units is then given by the expected output per craftsmen $\mathbb{E}(q_i)$ multiplied by the number of craftsmen N:

$$L = N\, k^\theta\, \Lambda.$$

In sum, all the results in the benchmark continue to hold with the "Leonardo" assumption, up to some constant terms: $\Gamma(1-\theta)$ should be replaced by Λ.

In the proof of Proposition 3, the "Leonardo" assumption requires to compute $\partial \mathbb{E}(q')/\, \partial k'$ differently. In particular, in deriving $\mathbb{E}(q')$ with respect to k' we should be careful in taking as given the future average society knowledge $\bar{q}' = (k')^\theta \Gamma(1-\theta)$, i.e. the externality in (3). More precisely, (3) should be written as:

$$\mathbb{E}(q') = \int_0^{\hat{h}'} \underbrace{(k')^\theta \Gamma(1-\theta)}_{\bar{q}\text{ (exogenous)}} k' \exp[-k' h_i] dh_i + \int_{\hat{h}'}^{\infty} h_i^{-\theta} k' \exp[-k' h_i] dh_i$$

where

$$\hat{h}' = \Gamma(1-\theta)^{-1/\theta}/k'.$$

Integrating we obtain:

$$\mathbb{E}(q') = (1 - \exp[\Gamma(1-\theta)/\theta])\, \bar{q} + (k')^\theta\, \Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})$$

and the derivative is

$$\frac{\partial \mathbb{E}(q')}{\partial k'} = \theta\Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})(k')^{\theta-1} = \Theta(k')^{\theta-1}.$$

with $\Theta = \theta\Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})$. Compared to the benchmark, there is a factor $\Gamma(1-\theta, \Gamma(1-\theta)^{-1/\theta})$ instead of a factor $\Gamma(1-\theta)$.

The model with the "Leonardo" assumption also speaks to the role of the unlimited support of knowledge in the baseline model. One may regard this assumption as unattractive, because it implies that everything that can be known is already known to at least someone from the beginning of time. The model extension shows that our results remain intact if the support of realized productivity is bounded. Moreover, one can regard this extension as an approximation of a model where latent knowledge is bounded as well, i.e., where there is a point mass of knowledge at the upper bound. The difference between such a model and the "Leonardo" extension is that in the "Leonardo" model the upper bound moves up in proportion to average knowledge, whereas with bounded support the upper bound would stay in place. Yet because growth of average knowledge is slow, the two models would be yield near-identical implications in the short and medium run (although if the rate of creation of new knowledge were zero, growth would peter out in the long run). Hence, one can regard our analysis as an approximation of a model with a finite support of knowledge.

C Proof of Proposition 2

The threshold for sufficient altruism is given by

$$\hat{\gamma} = \frac{(\delta(n^{\mathtt{F}}) - \kappa n^{\mathtt{F}})\,\bar{n}s}{(1-\alpha)(n^{\mathtt{F}})^2}\Gamma(1-\theta), \quad \text{with } n^{\mathtt{F}} = (1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}.$$

We claim that the balanced growth path exists if $\gamma > \max\{0, \hat{\gamma}\}$.

Let us first compute the balanced growth path, supposing it exists. The growth rate of knowledge in (b) comes from (13) where we have imposed $m = m^{\mathtt{F}} = 1$. Population growth is obtained using (10). Income per capita in (c) derives from (9). Utility in (d) is derived as follows: Income of a given craftsman is $q_i\,(1-\alpha)\,\frac{Y}{L} + \kappa a^{\mathtt{F}}$ (from (5)). Expected income, using (6), is $k^{\theta}\Gamma(1-\theta)(1-\alpha)\frac{Y}{L} + \kappa a^{\mathtt{F}} + \alpha y^{\mathtt{F}}$, where $\alpha y^{\mathtt{F}}$ is income from owning land. Given the value of L from (7), this simplifies into $(1-\alpha)\frac{Y}{N} + \kappa a^{\mathtt{F}} + \alpha y^{\mathtt{F}}$. The future labor income of the child is $(1-\alpha)y^{\mathtt{F}}$.

For the balanced growth path defined above to be incentive compatible, i.e. parents are indeed willing to provide their kids with teaching, the cost of teaching should be less than the gain in the income, as evaluated by the altruistic parents. Normalizing the labor income of a craftsman without training to zero, this condition is:

$$\delta(n^{\mathtt{F}}) - \kappa n^{\mathtt{F}} < \gamma\, n^{\mathtt{F}} E\left[(1-\alpha)q\frac{Y}{L}\right] = \gamma\, n^{\mathtt{F}}(1-\alpha)y^{\mathtt{F}}.$$

Using (6) and (7), this condition determines the lower bound on the altruism factor γ:

$$\hat{\gamma} > \frac{(\delta(n^{\mathtt{F}}) - \kappa n^{\mathtt{F}})\,\bar{n}s}{(1-\alpha)(n^{\mathtt{F}})^2}\Gamma(1-\theta).$$

D Proof of Proposition 3

The required threshold for altruism is given by:

$$\gamma > \frac{\delta'\left((1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}\right) - \kappa}{\frac{(1-\alpha)\nu\theta}{\bar{n}s}(1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}-2}}.$$

Let us first compute the balanced growth path, supposing it exists. To determine the number of apprentices per master, we use (14). Equation (15) is derived as follows. Each apprentice learns from $m^{\mathbf{C}} = (n^{\mathbf{C}})^o$ masters. As the draws for the initial knowledge of masters are not independent (all these masters were educated by the same persons), their acquired knowledge k^i is the same, but they had different ideas of their own drawn from $\mathrm{Exp}(\nu k_{-1})$. The acquired knowledge of the apprentices thus follows:

$$(k^i)' = k^i + m^{\mathbf{C}} \nu k_{-1}.$$

The final knowledge of the apprentices is given by

$$k' = (k^i)' + \nu k.$$

Using $(k^i)' = k' - \nu k$ and $(k^i) = k - \nu k_{-1}$ in the first equation, we get

$$k' - \nu k = k - \nu k_{-1} + m^{\mathbf{C}} \nu k_{-1} \tag{4}$$

which leads to (15) where $g = k'/k$.

The average utility expression takes into account the flows between master and apprentices. Adults are paid as masters by $a^{\mathbf{C}}$ parents the amount of $\delta'(a^{\mathbf{C}}) - \kappa$. Their disutility net of income from apprentices is $\delta(a^{\mathbf{C}}) - \kappa a^{\mathbf{C}}$. They also pay as parents the same amount $\delta'(a^{\mathbf{C}}) - \kappa$ for each of their children $n^{\mathbf{C}}$ to each of their master $m^{\mathbf{C}}$. The balance is:

$$a^{\mathbf{C}}(\delta'(a^{\mathbf{C}}) - \kappa) - (\delta(a^{\mathbf{C}}) - \kappa a^{\mathbf{C}}) - m^{\mathbf{C}} n^{\mathbf{C}} (\delta'(a^{\mathbf{C}}) - \kappa) = -(\delta(a^{\mathbf{C}}) - \kappa a^{\mathbf{C}}).$$

From Equation (15) and the value of $n^{\mathbf{C}}$, the growth rate $g = g^{\mathbf{C}}$ should satisfy:

$$g^2 - (1+\nu)g + \nu = \nu g^{\frac{(1-\alpha)\theta o}{\alpha}}. \tag{5}$$

The left hand side is a convex function $f_1(g)$, while the right hand side is a function of g and o, $f_2(g, o)$. Figure III represents these two functions for different sizes of the clan. At the minimum possible value of g, the left hand side is smaller than the right hand side:

$$f_1(1) = 0 < f_2(1, o) = \nu \forall o > 0.$$

Several cases may occur:

- For $o \leq \frac{\alpha}{(1-\alpha)\theta}$, $f_2(g, o)$ is concave, crosses $f_1(g)$ once, and there exist a solution to the equality (5).
- For $\frac{\alpha}{(1-\alpha)\theta} < o \leq \frac{2\alpha}{(1-\alpha)\theta}$, one can apply the l'Hospital rule twice to show that

  $$\lim_{g \to \infty} \frac{f_1(g)}{f_2(g, o)} > 1,$$

  implying that $f_2(g, o)$ is below $f_1(g)$ for large g and crosses it once. Hence there exist a solution to the equality (5).

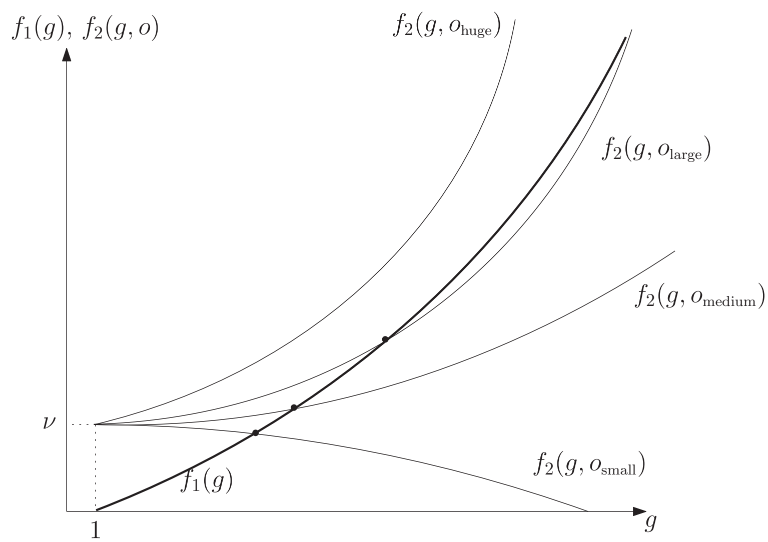

Figure III: Equation (5)

- If $o > \frac{2\alpha}{(1-\alpha)\theta}$ but not "too large," $f_2(g,o)$ cuts $f_1(g)$ twice. There are two balanced growth paths.
- For o very large, the function $f_2(g,o)$ which is above $f_1(g)$ for $g = 1$, stays above it as g increases. In that case, there is no solution to Equation (5), and no balanced growth path. The interpretation is that the clan is so big that technological level and population grow at an accelerating rate.

We can summarize these findings by defining $\bar{o}$, such that if $o \leq \bar{o}$, a balanced growth path exists.

Differentiating (5), the effect of o on g is given by:

$$\frac{dg}{do} = \frac{\partial f_2/\partial o}{\partial f_1/\partial g - \partial f_2/\partial g}.$$

The term $\partial f_2/\partial o$ is positive as $g > 1$. We conclude that increasing o increases g for equilibria where $\partial f_1/\partial g > \partial f_2/\partial g$, that is where $f_2(g,o)$ cuts $f_1(g)$ from above as g increases.

Let us now consider whether such an equilibrium is incentive compatible. If o is too low, the threat of punishment is insufficient to prevent shirking, and only the family equilibrium exists. If o is large, parents may no longer be willing to apprentice their children with all current adult members of the clan. In other words, the clan should be large enough for the threat of punishment to ensure compliance, but small enough for parents to be willing to pay the apprenticeship fee. Hence, there exists thresholds o_{min} and o_{max} such that if $o_{\text{max}} > o > o_{\text{min}}$ the clan equilibrium is incentive compatible.

Assuming that the punishment technology is such that one person can do only negligible harm to another, but more than one person can exert a much more damaging action, sufficient in any case to deter shirking, we get $o_{\min} = 0$.

The clan equilibrium is sustained if, for each child, the marginal cost paid to the master is lower than the expected marginal benefit, as priced by the parents:

$$\delta'(a) - \kappa \leq \gamma \frac{\partial \mathbb{E}\left(q'\right)}{\partial k'} \frac{\partial k'}{\partial m} \left(1-\alpha\right) \frac{Y'}{L'}. \tag{6}$$

The right hand side represents the expected effect on individual productivity of meeting an additional master. $(1-\alpha)\frac{Y'}{L'}$ is exogenous for the individual, and the altruism parameter γ reflects that the marginal benefit is seen from the point of view of the parent.

Let us first consider the term $\partial \mathbb{E}\left(q'\right)/\partial k'$. From (6), the derivative is

$$\frac{\partial \mathbb{E}\left(q'\right)}{\partial k'} = \theta \Gamma(1-\theta)(k')^{\theta-1}.$$

The term $\partial k'/\partial m$ can be directly obtained using (4) and is equal to νk_{-1}. Finally, the term Y'/L' can be transformed into:

$$\frac{Y'}{L'} = \frac{Y'}{N'}\frac{N'}{L'} = \frac{Y'}{N'}\frac{1}{\Gamma(1-\theta)(k')^{\theta}} \tag{7}$$

using (7). Condition (6) can now be rewritten as:

$$\delta'(a) - \kappa \leq \gamma\theta(k')^{\theta-1}\ \nu k_{-1}\ (1-\alpha)\frac{Y'}{N'}\frac{1}{(k')^{\theta}}. \tag{8}$$

which, along a balanced growth path, reduces to

$$\delta'\left((g^{\mathbf{C}})^{\frac{(1-\alpha)\theta(o+1)}{\alpha}}\right) - \kappa \leq \frac{\gamma(1-\alpha)\nu\theta}{\bar{n}s}\ (g^{\mathbf{C}})^{\frac{(1-\alpha)\theta}{\alpha}-2}.$$

The left hand side is increasing in o as $g^{\mathbf{C}} > 1$, $g^{\mathbf{C}}$ is increasing in o, and $\delta(a)$ is convex. If it is smaller than the right hand side for the minimum value of o ($o = 0$, $g^{\mathbf{C}} = 1+\nu$), i.e. if

$$\delta'\left((1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}}\right) - \kappa < \frac{\gamma(1-\alpha)\nu\theta}{\bar{n}s}(1+\nu)^{\frac{(1-\alpha)\theta}{\alpha}-2}.$$

then either the right hand side becomes larger than the left hand side for some value of $o = \hat{o} > 0$, or they never intersect, in which case $\hat{o}$ is infinite. Notice that the right hand side is decreasing in $g^{\mathbf{C}}$, and hence in o, provided that

$$(1-\alpha)\theta < 2\alpha,$$

in which case, $o_{\max}$, the maximum size of the clan which is incentive compatible, is necessarily finite.

In the analysis of the incentive compatibility, we have assumed that $g^{\mathbf{C}}$ was defined, we show above that it is not the case $o > \bar{o}$. Hence, the threshold above which a balanced growth path exists and is incentive compatible is $o_{\max} = \min\{\bar{o}, \hat{o}\}$.

E Proof of Lemma 4

The first-order condition for the parent's problem can be written as:

$$\underbrace{p}_{\text{marginal cost}} \quad n = \gamma\, n \underbrace{\frac{\partial \mathbb{E}\left(q'\right)}{\partial k'} \frac{\partial \mathbb{E}\left(k'\right)}{\partial m} \left(1-\alpha\right) \frac{Y'}{L'}}_{\text{marginal benefit } (\partial I'/\partial m)}.$$

Remembering from (6) that

$$\frac{\partial \mathbb{E}\left(q'\right)}{\partial k'} = \theta \Gamma (1-\theta)(k')^{\theta-1},$$

and using $k' = k(m+\nu)$ (from (13)) as masters are now drawn randomly, we obtain, in equilibrium:

$$\delta'(a) - \kappa = p = \gamma\theta\Gamma(1-\theta)\frac{(k')^{\theta}}{m+\nu}\,(1-\alpha)\frac{Y'}{L'}.$$

Using (7), this equation simplifies into

$$\delta'(a) - \kappa = \gamma\theta(1-\alpha)\frac{1}{m+\nu}\frac{Y'}{N'}.$$

F Proof of Proposition 5

For a fixed number of masters m, the market equilibrium yields higher growth in productivity because those masters are drawn randomly. Indeed, from the proof of Proposition 3, productivity in **C** follows

$$k' = (1+\nu)k + (m-1)\nu k_{-1}$$

which implies a growth rate equal to:

$$g^{\mathbf{C}} = \frac{1+\nu+\sqrt{(1+\nu)^2+4(m-1)\nu}}{2}$$

which is always less than the growth rate with the market, $m+\nu$, for $\nu > 0$ and $m > 1$.

Moreover, the equilibrium number of masters $m^{\mathbf{M}}$ is higher in the market equilibrium compared to the clan equilibrium. Indeed, $m^{\mathbf{M}}$ balances marginal cost and benefit. If the clan equilibrium had a higher number of masters, it would not be incentive compatible (parents would not like to pay all those masters).

Notice that, in the computation of the utility, the payments from apprentices, $pa^{\mathbf{M}}$, and for children, $pn^{\mathbf{M}}m^{\mathbf{M}}$, balance.

G Proof of Lemma 6

The maximization problem of the guild is:

$$\max_{a_j} \left\{ p_j a_j - \delta(a_j) + \kappa a_j \right\}$$

subject to:

$$(9) \qquad \begin{aligned} S_j N' m_j &= N a_j, \\ p_j &= \gamma \frac{\partial E I'_{ij}}{\partial m_j}, \\ \gamma E I'_{ij} - p_j m_j &= \gamma (1-\alpha) \frac{Y'}{N'} - pm. \end{aligned}$$

Replacing p_j and p by their value from the second constraint into the third leads to:

$$\gamma (S'_j)^{\frac{1}{\lambda}-1} \left(\frac{k'_j}{k'} \right)^{\frac{\theta}{\lambda}} - \frac{\theta}{m_j + \nu} (S'_j)^{\frac{1}{\lambda}-1} \left(\frac{k'_j}{k'} \right)^{\frac{\theta}{\lambda}} m_j = \gamma - \frac{\theta}{m+\nu} m.$$

which can be solved for S'_j:

$$S'_j = \left[\frac{(\gamma-\theta)\, m + \gamma\nu}{(\gamma-\theta)\, m_j + \gamma\nu} \left(\frac{m_j + \nu}{m+\nu} \right)^{1-\frac{\theta}{\lambda}} \right]^{\frac{\lambda}{1-\lambda}}.$$

The second constraint can be rewritten as:

$$p_j = \gamma\theta(1-\alpha) \frac{1}{m+\nu} \frac{Y'}{N'} \frac{(\gamma-\theta)\, m + \gamma\nu}{(\gamma-\theta)\, m_j + \gamma\nu}$$

and the equilibrium on the apprenticeship market (9) is:

$$a_j = n\, S'_j\, m_j.$$

These constraints imply that lowering the supply of apprenticeships a_j increases the price p_j. It is now easier to express the maximization program in terms of m_j:

$$\max_{m_j} \left\{ \left(\frac{\gamma\theta(1-\alpha)}{m+\nu} \frac{Y'}{N'} \frac{(\gamma-\theta)\, m + \gamma\nu}{(\gamma-\theta)\, m_j + \gamma\nu} + \kappa \right) n\, S'_j\, m_j - \delta\left(n\, S'_j\, m_j\right) \right\}.$$

The first order condition is:

$$\begin{aligned} &- (\gamma-\theta) \left(\frac{\gamma\theta(1-\alpha)}{m+\nu} \frac{Y'}{N'} \frac{(\gamma-\theta)\, m + \gamma\nu}{\left((\gamma-\theta)\, m_j + \gamma\nu\right)^2} \right) S'_j\, m_j + \\ &\left(\left(\frac{\gamma\theta(1-\alpha)}{m+\nu} \frac{Y'}{N'} \frac{(\gamma-\theta)\, m + \gamma\nu}{(\gamma-\theta)\, m_j + \gamma\nu} + \kappa \right) - \delta'\left(n\, S'_j\, m_j\right) \right) \left(\frac{\partial S'_j}{\partial m_j} m_j + S'_j \right) = 0 \end{aligned}$$

with

$$\frac{\partial S'_j}{\partial m_j} = \frac{\lambda}{1-\lambda}\left[\frac{(\gamma-\theta)\,m+\gamma\nu}{(\gamma-\theta)\,m_j+\gamma\nu}\left(\frac{m_j+\nu}{m+\nu}\right)^{1-\frac{\theta}{\lambda}}\right]^{\frac{\lambda}{1-\lambda}-1}$$
$$\times\left(-(\gamma-\theta)\frac{(\gamma-\theta)\,m+\gamma\nu}{\left((\gamma-\theta)\,m_j+\gamma\nu\right)^2}\left(\frac{m_j+\nu}{m+\nu}\right)^{1-\frac{\theta}{\lambda}}\right.$$
$$\left.+\left(1-\frac{\theta}{\lambda}\right)\frac{(\gamma-\theta)\,m+\gamma\nu}{(\gamma-\theta)\,m_j+\gamma\nu}\left(\frac{m_j+\nu}{m+\nu}\right)^{-\frac{\theta}{\lambda}}\right).$$

At the symmetric Nash equilibrium between guilds, this becomes:

$$-(\gamma-\theta)\left(\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}\,\frac{1}{(\gamma-\theta)\,m+\gamma\nu}\right)m+$$
$$\left(\left(\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}+\kappa\right)-\delta'\left(n\,m\right)\right)\left(\frac{\partial S'_j}{\partial m_j}\,m+1\right)=0$$

with:

$$\frac{\partial S'_j}{\partial m_j} = \frac{\lambda}{1-\lambda}\left(\frac{-(\gamma-\theta)}{(\gamma-\theta)\,m+\gamma\nu}+\left(1-\frac{\theta}{\lambda}\right)\right),$$

which we can rearrange into:

$$\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}\left(\frac{-(\gamma-\theta)\,m}{(\gamma-\theta)\,m+\gamma\nu}+\frac{\partial S'_j}{\partial m_j}\,m+1\right)+(\kappa-\delta'\left(n\,m\right))\left(\frac{\partial S'_j}{\partial m_j}\,m+1\right)=0$$

and finally:

$$\delta'(nm)-\kappa = \left(\frac{\gamma\theta(1-\alpha)}{m+\nu}\,\frac{Y'}{N'}\right)\Omega(m)$$

with:

$$\Omega(m) = \frac{\left(\frac{-(\gamma-\theta)m}{(\gamma-\theta)m+\gamma\nu}+1+\frac{\lambda m}{1-\lambda}\left(\frac{-(\gamma-\theta)}{(\gamma-\theta)m+\gamma\nu}+\left(1-\frac{\theta}{\lambda}\right)\right)\right)}{1+\frac{\lambda m}{1-\lambda}\left(\frac{-(\gamma-\theta)}{(\gamma-\theta)m+\gamma\nu}+\left(1-\frac{\theta}{\lambda}\right)\right)}.$$

$\Omega(m)$ can be further simplified into:

$$\Omega(m) = \frac{1-\lambda+(\lambda-\theta)m-\frac{(\gamma-\theta)m}{(\gamma-\theta)m+\gamma\nu}}{1-\lambda+(\lambda-\theta)m-\lambda\,\frac{(\gamma-\theta)m}{(\gamma-\theta)m+\gamma\nu}}.$$

When $\lambda\to 1$, $\frac{\partial S'_j}{\partial m_j}\to\infty$ and $\Omega(m)\to 1$.

H Proof of Proposition 7

Before considering the optimization problem, let us compute the effect of changing m_j on income I_j. Using Equation (7), we can compute the relative quantity of efficient labor in sector j as:

$$\frac{L'_j}{L'} = S'_j \left(\frac{k'_j}{k'}\right)^{\theta}.$$

With this expression, and with Equation (6), which adapts to trade j as $Eq_j = \Gamma(1-\theta)k_j^{\theta}$, it is convenient to rewrite expected income as:

$$EI'_{ij} = (1-\alpha)\,\frac{Y'}{N'}(S'_j)^{\frac{1}{\lambda}-1}\left(\frac{k'_j}{k'}\right)^{\frac{\theta}{\lambda}}.$$

Let us now compute the effect of changing m_j on individual income. Using the result in (7), and $k'_j = (m_j + \nu)k_j$, we get:

$$\frac{\partial EI'_{ij}}{\partial m_j} = \theta\Gamma(1-\theta)\frac{(k'_j)^{\theta}}{m_j+\nu}\,(1-\alpha)\,\frac{Y'}{L'}\left(\frac{L'_j}{L'}\right)^{\frac{1}{\lambda}-1}$$

which can moreover be simplified using (7):

$$\frac{\partial EI'_{ij}}{\partial m_j} = \theta(1-\alpha)\frac{(k'_j)^{\theta}}{(k')^{\theta}}\frac{1}{m_j+\nu}\,(1-\alpha),$$

leading to:

$$\frac{\partial EI'_{ij}}{\partial m_j} = \theta(1-\alpha)\frac{1}{m_j+\nu}\,\frac{Y'}{N'}(S'_j)^{\frac{1}{\lambda}-1}\left(\frac{k'_j}{k'}\right)^{\frac{\theta}{\lambda}}.$$

Marginal income is therefore equal to expected income multiplied by: $\frac{\theta}{m_j+\nu}$.

To study the equilibrium, one should consider Equation (24) replacing $a^{\mathbf{G}}$, $m^{\mathbf{G}}$ and $y^{\mathbf{G}}$ by their value:

$$\delta'((g^{\mathbf{G}}-\nu)n^{\mathbf{G}}) - \kappa = \gamma\theta(1-\alpha)\frac{1}{g^{\mathbf{G}}}\;\frac{n^{\mathbf{G}}}{\bar{n}s}\;\Omega(g^{\mathbf{G}}-\nu). \tag{10}$$

This equation describes a relationship between $g^{\mathbf{G}}$ and $n^{\mathbf{G}}$ which we call the "apprenticeship monopolistic market," as it is derived from the demand for apprenticeship and the monopolistic behavior of the guild. Equation (10) can be rewritten as:

$$n^{\mathbf{G}} = \frac{\kappa}{\bar{\delta}(g^{\mathbf{G}}-\nu) - \dfrac{\gamma\theta(1-\alpha)}{s\bar{n}\,g^{\mathbf{G}}}\Omega(g^{\mathbf{G}}-\nu)}.$$

If we compare this expression with the equivalent in the market equilibrium, Equation (20), we see that the denominator is necessarily larger. Hence, for any given g, $n^{\mathbf{G}} < n^{\mathbf{M}}$. It follows that $g^{\mathbf{G}} < g^{\mathbf{M}}$.

Notice finally that, as in the market equilibrium, the payments from apprentices, $pa^{\mathbf{G}}$, and for children, $pn^{\mathbf{G}}m^{\mathbf{G}}$, balance in the computation of the utility.

I Proof of Proposition 8

We can compute the gains of adopting the guild institution as:

$$\begin{aligned} u^{\mathbf{F}\to\mathbf{G}} - u^{\mathbf{F}\to\mathbf{F}} &= \gamma n_0(y_1^{\mathbf{G}} - y_1^{\mathbf{F}}) + \kappa(a_0 - n_0) - \delta(a_0) + \delta(n_0) - \mu(N_0), \\ u^{\mathbf{C}\to\mathbf{G}} - u^{\mathbf{C}\to\mathbf{C}} &= \gamma n_0(y_1^{\mathbf{G}} - y_1^{\mathbf{C}}) + \kappa(a_0 - (n_0)^{o+1}) - \delta(a_0) + \delta((n_0)^{o+1}) \\ &\quad -\mu(N_0). \end{aligned}$$

$\underline{N}$ makes people in the family equilibrium indifferent between adopting the guild or not, i.e., it solves:

$$\gamma n_0(y_1^{\mathbf{G}} - y_1^{\mathbf{F}}) + \kappa(a_0 - n_0) - \delta(a_0) + \delta(n_0) - \mu(\underline{N}) = 0.$$

One should show that for $N_0 = \underline{N}$, people in the clan equilibrium do not want to adopt the guild, i.e.:

$$\gamma n_0(y_1^{\mathbf{G}} - y_1^{\mathbf{C}}) + \kappa(a_0 - (n_0)^{o+1}) - \delta(a_0) + \delta((n_0)^{o+1}) - \mu(\underline{N}) < 0.$$

This is true if:

$$\gamma n_0 y_1^{\mathbf{C}} + (\kappa(n_0)^{o+1} - \delta((n_0)^{o+1})) > \gamma y_1^{\mathbf{F}} + \kappa n_0 - \delta(n_0). \tag{11}$$

Let us define the following function:

$$\psi(m) = \gamma n_0 u(m) + \kappa m n_0 - \delta(m n_0).$$

$u(m)$ is the function that relates future income to number of masters learning from, in the context of the clan equilibrium. From (8) and (4), we get:

$$u(m) = \Gamma(1-\theta)^{1-\alpha} (\underbrace{(1+\nu)k - \nu k_{-1} + m\nu k_{-1}}_{k'})^{(1-\alpha)\theta} (N')^{-\alpha}.$$

Hence, $u(m)$ is increasing and concave in m. As a consequence, $\psi(m)$ is also concave in m ($\delta(\cdot)$ is convex). To see whether it is increasing, we can compute:

$$\psi'(m) = \gamma\, n_0(1-\alpha)\Gamma(1-\theta)^{1-\alpha}\theta(k')^{(1-\alpha)\theta-1}\nu k_{-1}\,(N')^{-\alpha} + \kappa n_0 - \delta'(m n_0) n_0.$$

We also know from Appendix D that the clan equilibrium is sustained if the marginal cost paid to the master is less than the expected marginal benefit, i.e. $\delta'((n_0)^{o+1}) - \kappa \leq \partial y_1^{\mathbf{C}}/\partial m_0$, which implies, from (8) and (8):

$$\gamma\theta\,(1-\alpha)\Gamma(1-\theta)^{1-\alpha}\,(k')^{(1-\alpha)\theta-1}\nu k_{-1}\,(N')^{-\alpha} + \kappa - \delta'(m n_0) \geq 0.$$

This individual level condition implies that, at the aggregate equilibrium, $\psi'(m) > 0$. Using the mean value theorem for derivatives, we know there exists $\tilde{m} \in [1, m^{\mathbf{C}}]$ such that $(\psi(m^{\mathbf{C}}) - \psi(1))/(m^{\mathbf{C}} - 1) = \psi'(\tilde{m})$. As $\psi(\cdot)$ is concave, $\psi'(\tilde{m}) > \psi'(m^{\mathbf{C}}) > 0$ which proves $\psi(m^{\mathbf{C}}) > \psi(1)$ and inequality (11) holds.

$\overline{N}$ makes people in the clan equilibrium indifferent between adopting the guild or not, i.e. it solves

$$\gamma n_0(y_1^{\mathbf{G}} - y_1^{\mathbf{C}}) + \kappa(a_0 - (n_0)^{o+1}) - \delta(a_0) + \delta((n_0)^{o+1}) - \mu(\overline{N}) = 0.$$

One should show that for $N_0 = \overline{N}$, people in the family equilibrium also want to adopt the guild, i.e.:

$$\gamma n_0(y_1^{\mathbf{G}} - y_1^{\mathbf{F}}) + \kappa(a_0 - n_0) - \delta(a_0) + \delta(n_0) - \mu(\overline{N}) > 0.$$

This is true as $\mu(\overline{N}) < \mu(\underline{N})$.

J Productivity Growth Calculations for Western Europe and China

Table I displays the data that underlie our calculations of aggregate productivity growth in Western Europe and China in the period 1–1820. We compute productivity based on the aggregate production function used in the main analysis with a TFP term A_t:

$$Y_t = A_t P_t^{1-\alpha} X^{\alpha},$$

where $X = 1$ is the fixed factor land (the normalization to one amounts to a choice of units) and P_t is labor supply, measured by population. Productivity can therefore be computed from data on population and GDP per capita (Y_t/P_t) as:

$$A_t = \frac{Y_t}{P_t^{1-\alpha}} = \left(\frac{Y_t}{P_t}\right) p_t^{\alpha}. \tag{12}$$

As discussed in Section V.A., the share of land is set to $\alpha = 1/3$. We do not explicitly account for physical capital, which was a relatively unimportant factor of production in the pre-industrial period, and at any rate no reliable long-term estimates of capital are available. Any changes in output that are due to physical investment would be reflected in measured productivity.

The first two panels of Table I display population and income levels Western Europe and China from Maddison (2010). From 1 to 1000, the data imply slowly growing TFP in China and actually some regression in Western Europe. After 1000, there is an acceleration in productivity growth in both regions, but the rise in much more pronounced in Europe, with a gap between the regions of 2.5 percent per period until 1500, and 2.4 percent from 1500 to 1820.

Given the distant time periods involved, estimating income levels is naturally fraught with difficulty. As an alternative data source, we also consider the recent estimates of levels of income per capita by Broadberry, Guan, and Li (2014). The numbers are not always directly comparable to Maddison (2010) (e.g., Broadberry, Guan, and Li 2014 do not provide an average for Western Europe). The sources broadly agree on the measurement of population and on the acceleration of productivity growth in Europe, but there are large deviations in the estimates of levels of income per capita in China, which matter for the computation of growth rates. To generate an alternative estimate of the gap in productivity growth between Western Europe and China after 1000, in the third panel of Table I we combine the population data from Maddison (2010)

Year	Population	GDP/cap.	TFP	ΔTFP
Western Europe (Maddison 2010)				
1	18,600	600	159	
1000	19,700	425	115	-0.8%
1500	48,192	797	290	4.7%
1820	114,559	1,234	599	5.8%
China (Maddison 2010)				
1	59,600	450	176	
1000	59,000	466	181	0.1%
1500	103,000	600	281	2.2%
1820	381,000	600	435	3.5%
China (GDP/cap.: Broadberry et al. 2014)				
1000	59,000	1,382	538	
1500	103,000	1,127	528	-0.1%
1820	381,000	595	431	-1.6%

TABLE I: Growth Accounting for China and Western Europe, 1–1820. Population in 1,000, income per capita in 1990 international dollars, TFP is computed according to (12) and divided by 100, ΔTFP is growth rate of TFP in percent per generation (25 years).

with the estimates of GDP per capita from Broadberry, Guan, and Li (2014). The estimates for GDP per capita in China in 1000 and 1500 (the year 1 is not available) are much higher compared to Maddison (2010), resulting in a decline in measured TFP over time. Based on these numbers, the gap in TFP growth between Western Europe and China per 25 years was 4.8 percent for 1000–1500, and 7.4 percent for 1500–1820.

A caveat in using these gaps in overall TFP growth to evaluate our model is that our model is primarily about productivity improvements in artisanal production, whereas agriculture made up the majority of the economy. This can be partly addressed by noting that there are important interactions between artisans and craftsmen and agricultural productivity, such as improvements driven by better agricultural implements, dairying techniques, millwrighting, and better storage, transport, and processing of agricultural products, which concerns a large number of crafts. In addition, some of the benefits of better dissemination of knowledge in Europe can be found in agriculture in altered form. One aspect of this is the prevalence in Western Europe of highly mobile servants, both male and female, who worked on farms other than those of their immediately family and may have contributed to the dissemination of productivity improvements. The prevalence of mobile servants in Western Europe was due to differences in family organization around the world (see Hajnal 1982) and did not have a close equivalent in China. Regarding other productivity improvements in agriculture, the key issue for our purposes is whether these were much different between China and Western Europe, i.e., whether the rise of Europe could perhaps be due primarily due to agricultural productivity rather than productivity in the crafts and ultimately manufacturing. The data to settle this question definitively does not exist, but our reading of the existing evidence is that at the very least improvements in the sectors that we focus on here played a very important role. Within agriculture, the main drivers of productivity were new crops and better cultivation methods. With regards to crops, an important factor was the introduction of new crops from the New World after Columbus, which occurred in both Europe and China. With regards to cultivation, Europe gained from the introduction of fodder crops and the closer integration of the pastoral and arable sectors in the "new husbandry," and China gained from the spread of double-cropping. What we think is beyond any question is the well-documented growth in the productivity of pre-Industrial Revolution (artisanal) manufacturing in Europe in a wide range of industries, from high-skilled instrument and clock making to ironworking and textiles. This was simply not matched by anything we see in China.

References

Broadberry, Stephen, Hanhui Guan, and David D Li. 2014. "China, Europe and the Great Divergence: A Study in Historical National Accounting, 980-1850." Economic History Department Paper, London School of Economics.

Hajnal, John. 1982. "Two Kinds of Preindustrial Household Formation System." *Population and Development Review*, pp. 449–94.

Maddison, Angus. 2010. "Statistics on World Population, GDP and Per Capita GDP, 1–2008 AD." Available from http://www.ggdc.net/maddison/oriindex.htm.

Mokyr, Joel. 2002. *The Gifts of Athena. Historical Origins of the Knowledge Economy.* Princeton: Princeton University Press.

Vollrath, Dietrich. 2011. "The Agricultural Basis of Comparative Development." *Journal of Economic Growth* 16 (4): 343–370.